The Complex Interplay between Multisensory Integration and Perceptual Awareness

Deroy, O.; Faivre, N.; Lunghi, C.; Spence, C.; Aller, M.; Noppeney, U.

DOI:
10.1163/22134808-00002529

Document Version
Peer reviewed version

Citation for published version (Harvard):
Deroy, O, Faivre, N, Lunghi, C, Spence, C, Aller, M & Noppeney, U 2016, 'The Complex Interplay between Multisensory Integration and Perceptual Awareness', *Multisensory Research*, vol. 29, no. 6, pp. 585-606. https://doi.org/10.1163/22134808-00002529

The complex interplay between multisensory integration and perceptual awareness

O. Deroy[1]*, N. Faivre[2]*, C. Lunghi[3]*, C. Spence[4]*, M. Aller[5], & U. Noppeney[5]

* Equal contributions, alphabetical order

[1]Centre for the Study of the Senses, Institute of Philosophy, School of Advanced Study, University of London, London, United Kingdom

[2]Laboratory of Cognitive Neuroscience, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne, Switzerland

[3]Department of Translational Research on New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy

[4]Crossmodal Research Laboratory, Department of Experimental Psychology, Oxford University, Oxford, United Kingdom

[5]Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, Birmingham, United Kingdom

Corresponding author:

Máté Aller

Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

Email: MXA342@student.bham.ac.uk

Keywords:

multisensory integration, perceptual awareness, consciousness, perception, crossmodal integration, metacognition

Abstract

The integration of information has been considered a hallmark of human consciousness, as it requires information being globally available via widespread neural interactions. Yet, the complex interdependencies between multisensory integration and perceptual awareness, or consciousness, remain to be defined. While perceptual awareness has traditionally been studied in a single sense, in recent years we have witnessed a surge of interest in the role of multisensory integration in perceptual awareness.

Based on a recent IMRF symposium on multisensory awareness, this review discusses three key questions from conceptual, methodological and experimental perspectives:

1. What do we study when we study multisensory awareness?

2. What is the relationship between multisensory integration and perceptual awareness?

3. Which experimental approaches are most promising to characterize multisensory awareness?

We hope that this review paper will provoke lively discussions, novel experiments, and conceptual considerations to advance our understanding of the multifaceted interplay between multisensory integration and consciousness.

Introduction

In our everyday lives, our brain has to deal with a constant influx of sensory signals. Looking at perceptual experience though, a fundamental aspect of our conscious awareness is that sensory signals are integrated nearly effortlessly into a seamless multisensory perception of our environment. Yet, even though multisensory experience is pervasive in everyday life, the relationship between multisensory integration and perceptual awareness remains unclear.

This lack of clarity is all the more surprising given that several leading theories see a strong link between information integration and perceptual awareness. For instance, according to the global workspace model, consciousness emerges when information is made globally available via long-range connectivity such as the fronto-parietal system (Dehaene 2001). Other theories suggest that consciousness emerges via recurrent interactions that enable information exchange across multiple levels of the cortical hierarchy (Lamme 2006; Lamme & Roelfsema 2000). Finally, the integrated information theory of consciousness associates consciousness with 'integrated information' and aims to determine the structural and functional properties that enable neural systems to form complex integrated information as a prerequisite of consciousness (Balduzzi & Tononi 2008).

Yet, despite the proposed link between information integration and consciousness, perceptual awareness has traditionally been studied in terms of single-sense experiences (see De Graaf et al. 2012; Dehaene & Changeux 2011 for reviews), such as vision, audition (Allen et al. 2000; Bekinschtein et al. 2009; Giani et al. 2015; Gutschalk et al. 2008; Haynes et al. 2005; Ro et al. 2003), or, on occasion, touch (Gallace & Spence 2008; Gallace & Spence 2014) or olfaction (Stevenson & Attuquayefio 2013). Only in recent years have we witnessed a surge of interest in studying perceptual awareness in multisensory terms. Based on a recent IMRF symposium on the topic, this paper aims to review the key conceptual, methodological and empirical findings that have advanced the field in recent years, and to provide better tools to

confront the challenges raised by the multifaceted interplay between multisensory integration and perceptual awareness.

Section 1 provides a conceptual map of the kind of phenomena which fall under the general label of 'multisensory awareness', and highlights the main challenge for the field. We discuss the commonalities and differences of perceptual awareness that may occur in unisensory and multisensory contexts. For instance, in vision, information needs to be integrated across time and space into a coherent percept of our dynamic environment. Vision also faces the challenge of binding features such as colour and form, which are represented predominantly in different brain areas, into a unified object percept (Ghose & Maunsell 1999; Roskies 1999; Wolfe & Cave 1999). Along similar lines, multisensory perception relies on binding complementary pieces of information (e.g. an object's shape from the front side via vision and from the rear via touch) that are provided by different sensory modalities. Moreover, different senses can provide redundant information about specific properties such as the spatial location or timing of an event.

Section 2 explores the relationship between multisensory integration and multisensory awareness. More specifically, it reviews the behavioural and neural research investigating the extent to which multisensory signals can be integrated in the absence of awareness. Numerous studies have demonstrated that signals that we are aware of in one sensory modality can boost signals from another sensory modality that we are not aware of into perceptual awareness depending on temporal coincidence, spatial or higher order correspondences such as semantic or phonological congruency (Adam & Noppeney 2014; Aller et al. 2015; Alsius & Munhall 2013; Chen & Spence 2011a; Chen & Spence 2011b; Hsiao et al. 2012; Olivers & Van der Burg 2008; Palmer & Ramsey 2012). Less is known about whether signals that we are unaware of can also influence where and how we perceive signals that we are aware of. Moreover, despite the vast neurophysiological evidence showing

multisensory interactions in anaesthetized animals (Stein & Meredith 1993) only little behavioural evidence has been accumulated indicating that two signals from different sensory modalities can interact in the absence of awareness such as sleep (Arzi et al. 2012) or when signals are masked and thus precluded from awareness in both sensory modalities (Faivre et al. 2014).

Finally, Section 3 discusses the various experimental approaches that can be pursued to tap into multisensory awareness. Unisensory research has developed a large repertoire of experimental manipulations and paradigms to contrast sensory processing in the presence and absence of awareness including multistable perception (e.g., ambiguous figures, multistable motion quartets, binocular rivalry, and continuous flash suppression), attentional blink, masking, or sleep. Which of those experimental approaches might be most promising when it comes to multisensory awareness?

Section 1: What do we study when we study multisensory awareness?

Most of our conscious experiences occur in a multisensory setting when several sensory modalities are likely being stimulated simultaneously. Some senses, like the vestibular system, proprioception, or touch, indeed almost never 'switch off' in natural circumstances. Meanwhile, audition and vision often function together starting with saccadic coordination (Heffner & Heffner 1992a, b; Kruger et al. 2014) and leading to many well-known audio-visual illusions, such as the spatial ventriloquist effect (Alais & Burr 2004; Bertelson & Aschersleben 1998; Vroomen & de Gelder 2004), the McGurk effect (McGurk & MacDonald 1976), the double-flash fission or fusion illusion (Andersen et al. 2004; Shams et al. 2000), and pitch-induced illusory motion (Maeda et al. 2004).

Most phenomenological reports also tell us that conscious experiences are multisensory: We perceive talking faces, we go through scented and colourful gardens, filled with birdsong, we sense the noise and feel of the computer keys pressed under our fingers. The evidence, then,

converges in making consciousness a matter of multisensory combination. This raises an important question: How should we map the concept of multisensory integration with the first-person evidence of unified perceptual awareness? It is important to note that integration is studied as a process, or rather a set of processes, while consciousness is often analysed as a state presenting us with objects, events, and their relations. With consciousness being one of the most discussed and controversial notion in the philosophical and scientific literature, we only attempt here to provide a useful taxonomy to distinguish between different cases of multisensory awareness, for the field to study. With these distinctions in hand, it is useful to look at what the study of awareness really involves, by drawing on two useful conceptual distinctions between access and content.

Three kinds of multisensory contents

While the field is most concerned with cases where a single property is perceived through two or more sensory modalities, there is more to multisensory awareness than these. The most studied cases concern those situations where different senses provide redundant information about specific properties such as the spatial location or timing of an event. Imagine, for instance, running through the forest and spotting a robin sitting on the branch and singing (Rohe & Noppeney 2015, 2016). By integrating redundant spatial information from vision and audition, the brain can form more reliable estimates of the location of the singing bird. Redundant information can even be provided about higher-order aspects such as a phoneme, as in speech perception. In fact, perceptual illusions such as spatial ventriloquism (Alais & Burr 2004; Bertelson & Aschersleben 1998; Vroomen & de Gelder 2004) or the McGurk illusion (Gau & Noppeney 2016; McGurk & MacDonald 1976; Munhall et al. 1996) emerge because different sensory modalities provide redundant, yet slightly conflicting information about the spatial location or a particular phoneme (e.g., [ba] vs. [ga]). Another

good illustration can be when touch and vision contribute to the perception of shape (Ernst & Banks 2002). This said, the integration of redundant information is not necessarily tied to multisensory awareness, and could lead to episodes of unisensory consciousness being biased by the information provided by another modality, be it consciously perceived or not. In other words, many of these cases could be cases of **crossmodal bias of unisensory awareness**, as much as genuine cases of unified multisensory awareness. A possible way to exclude the first possibility is to show that the integration of two sensory inputs leads to the conscious experience of a new property or aspect which could not be experienced by a conjunction of unisensory episodes. This could consist in being able to experience the simultaneity between two unisensory events, or in the emergence of a new quality, such as flavour, which is commonly taken to involve a fusion of taste, smell (retronasal olfactory), and trigeminal inputs (Spence et al. 2015). In other words, we should not be too fast in thinking that all cases of integration need to get manifested in episodes of multisensory awareness and should look for evidence of specific or **emerging multisensory properties** (see Partan & Marler 1999).

At least two other kinds of cases also need to be considered when studying multisensory awareness, besides cases resting on the integration of redundant information.

On the one hand, two modalities can contribute to the perception of the same object, but different or complementary properties of that object. These are the classical cases of multisensory binding- cases where one is conscious of the visual shape of the dog and the sound of its bark (Chen & Spence 2010), the shape of the kettle and the whistling sound (Jackson 1953). The two unisensory components need to be referred to the same object, or at least, in the case of an event, to the same moment and perceived location. These cases form a distinct category of conscious perception of **multisensory objects and events**, and raise different challenges than the one where the contents experienced by two modalities are the

same. Here the two contents can remain unisensory but multisensory awareness seems to be of their co-attribution to the same object (or space/time).

On the other hand, cases where different senses contribute to the perception of one and the same property, or object, should not make us forget about a third, and no less important, category where two objects in different modalities, or even two multisensory objects, are experienced as part of the same multisensory scene. For instance, you may be aware of the cup in front of you, while also being conscious of the shape and temperature of the spoon in your hand, and the sounds of the barista talking behind you. All these various unisensory and multisensory objects or events are different and yet they are all experienced as part of the same setting or scene - presenting us with a third kind of multisensory awareness, i.e. **multisensory scene perception**[1].

Multisensory contents vs. multisensory access

With these distinctions in hand, we can now turn to another important conceptual difference between what people report and what they are phenomenally aware of, or what they attend to. While there is no doubt that people will report experiencing multisensory objects, for instance, or scenes, the question that cognitive neuroscientists need to ask is whether these correspond to what is present in consciousness at any given time, or what is reconstructed through other processes aggregating information experienced at different times. If this crucially marks the difference between genuine multisensory awareness and other processes where conscious information can be coalesced, the difference is certainly easier to draw conceptually than experimentally. Spence & Bayne 2015, for instance, question whether reports of multisensory events or objects should be taken at face value for being about a unified conscious episode, and whether they do not perhaps hide a rapid switch of attention

[1] This aspect is seldom approached in the experimental literature, and more often discussed in the philosophical literature as a form of 'phenomenal unity'. See Deroy (2014) for review and discussion.

between unisensory conscious episodes. The co-attribution to a single object would then not depend on the awareness of a multisensory object but on something like an ‘attentional glue’. In the absence of a good model for how attention could perform this role, it might be sufficient to note that the co-attribution to a single object might be a matter of non-conscious representation, which keeps track of, and predicts a relation of co-occurrence and co-localisation between two properties (Deroy 2014). In other words, researchers interested in the link between multisensory integration and awareness should not take for granted that the kind of contents described above and reported by participants require multisensory access (see Table 1). Room should be left to explore how contents and access could come apart. While there is good evidence that we keep track of multisensory contents, a key question is to know whether those get manifested in consciousness or sit outside awareness. A second key question will be to see whether the same process or analysis should be given for all these cases. Integrating redundant information across the senses on the assumption that they concern a single property, or having to determine whether two kinds of information need to be referred to the same object, or how objects then relate to one another in a scene are different processes; It is likely that each will require to be investigated separately when it comes to its dependence on, and manifestation in, consciousness.

	Integration of information regarding a single property	Attribution of different properties referred to the same object	Co-presence of multiple objects in the same scene
Content	Multisensory property	Multisensory object	Multisensory scene
Access	Being aware of a single property across different senses at the same time	Being aware that two unisensory properties belong to the same object or are part of a single event	Being aware that two unisensory or multisensory objects are present at the same time in the environment

Table 1: Overview of the three kinds of cases falling under the heading of multisensory awareness. Evidence of multisensory access is different from evidence that our brains and minds are integrating information about properties, objects and scenes, as these contents could be the result of unconscious processes, and not experienced at once.

Section 2: What is the relationship between multisensory integration and awareness?

One of the key functions of the human brain is to monitor bodily states (interoception) and environmental states (exteroception) (Blanke 2012; Critchley & Harrison 2013; Faivre et al. 2015). Despite the tremendous amount and variability of exteroceptive and interoceptive signals the brain has to process, such monitoring seems to be performed flawlessly, and one experiences being an integrated bodily-self, evolving in a unified, multisensory world (i.e., phenomenal unity, Chalmers & Bayne, 2003). Intuitively, perceptual consciousness, (i.e., the subjective experiences caused by a subset of perceptual processes), may be better characterized as multisensory by essence, reflecting multisensory wholes rather than sums of unisensory features. In this respect, it is important to distinguish situations in which percepts from different modalities merely coexist (e.g., reading while scratching my hand), with

situations in which they merge into a single unitary experience (e.g., looking at my hand being scratched; Deroy et al. 2014). Many theories of perceptual consciousness postulate strong interdependencies between consciousness and the capacity to integrate information across the senses, but also across spatial, temporal, and semantic dimensions (Mudrik et al. 2014). Accordingly, when consciously processing signals of multiple sensory origins, one may have privileged access to the integrated product while losing access to its component parts, and therefore experience phenomenal unity. Exploring the properties of phenomenal unity empirically is challenging, considering the nonspecific nature of subjective report ("Did you experience a multisensory object or two unisensory features?"), but also the discrepancy between phenomenal experience and multisensory integration as measured at the neural level (Deroy et al. 2014). Initial evidence has shown that participants do not integrate signals from vision and haptics into perceptual metamers, but were still able to distinguish between perceptual estimates based on congruent and incongruent signals (Hillis et al. 2002). These results suggest that participants had at least to some extent access to the sensory component signals rather than one unified multisensory estimate.

At the behavioural level, it has been repeatedly shown that the processing of an invisible stimulus is affected by the processing of supraliminal stimuli in the auditory, tactile, proprioceptive, vestibular, or olfactory modalities (see below for details). Yet because in these studies participants were always conscious of the non-visual stimulus, these results could well reflect the interplay between unconscious vision and conscious processes in another modality, rather than an integrative process between two unconscious representations. Information about the supraliminal stimulus is possibly broadcast throughout the brain, and modulates visual neurons activated by the invisible stimulus.

Thus, these results are compatible with the view that multisensory integration requires consciousness, but we will now see that other studies in which no stimulus is consciously

perceived are more decisive. In one of them (Arzi et al. 2012), it was shown that associations between tones and odours occurred during NREM sleep, arguably in the complete absence of awareness. The authors relied on partial-reinforcement trace conditioning, and measured sniff responses to tones previously paired with pleasant and unpleasant odours while participants were sleeping. Even though subjects were in NREM sleep stage, and arguably unconscious, they sniffed in response to tones alone, suggesting that they learned novel multimodal associations unconsciously. However, controlling stimulus awareness during sleep is difficult, and the possibility remains that the stimuli were consciously accessed when presented, but forgotten by the time of awaking. In another study trying to account for this potential limitation, awake participants were shown to compare the numerical information conveyed by an invisible image and an inaudible sound (Faivre et al. 2014). Interestingly, such unconscious audiovisual comparisons only occurred in those cases where the participants had previously been trained with consciously perceived stimuli, thus suggesting that conscious but not unconscious training enabled subsequent unconscious processes. The level at which the comparison of written and spoken digits operates is still an open question. While it could involve multisensory analyses of low-level visual and acoustic features, a possibility remains that the comparison is made independently of perceptual features, once the visual and auditory stimuli have separately reached an amodal, semantic representation. Moreover, multisensory comparisons (e.g., congruency judgments) do not necessarily imply multisensory integration. Future studies may potentially help in disentangling these various mechanisms (Noel et al. 2015). First, disrupting the spatiotemporal structure of the audiovisual stream should have a larger impact on the comparative process in the case that it operates at a perceptual, rather than semantic, non-perceptual level. Second, in case the results rely on multisensory interactions in the absence of awareness, weakening the visual and auditory signals may potentially increase the strength of their integrated product, by

virtue of audiovisual inverse effectiveness (Stanford 2005; Stein et al. 2009; von Saldern & Noppeney 2013). Third, if two subliminal signals are indeed integrated into a unified percept rather than only compared, the integrated percept should be able to prime subsequent perceptual processing.

At the neural level, several mechanisms support the possibility of unconscious multisensory integration. First, unconscious multisensory integration may be enabled by multisensory neurons that that do not take part in large-scale interactions. While such neurons have been described at relatively low-levels in the brain, including primary sensory (Ghazanfar & Schroeder 2006; Kayser et al. 2010; Lee & Noppeney 2014; Lee & Noppeney 2011; Liang et al. 2013; Meyer et al. 2010; Rohe & Noppeney 2016; Vetter et al. 2014; Werner & Noppeney 2010a) and subcortical structures like the superior colliculus (see Meredith & Stein 1986; Stein & Stanford 2008, for a review), their relevance for elaborate cognitive functions remains to be assessed. Second, and higher in the neural hierarchy, another possibility is that unconscious multisensory integration operates through feedforward connections (and most likely outside of awareness, see Lamme & Roelfsema, 2000) between sensory cortical areas and multisensory convergence zones such as the superior temporal sulcus or the posterior parietal cortex (Schroeder & Foxe 2005; Werner & Noppeney 2010b). Interestingly, such feedforward processes within low-level cortices and at early post-stimulus latencies have been shown to affect multisensory information processing and behaviour despite stimulus unawareness (e.g., phosphene perception enhancement by unconscious looming sounds, Romei et al. 2009). In contrast with these mechanisms, multisensory integration is sometimes held to require long-range feedback connections between sensory cortices and frontoparietal networks, a mechanism that typically coincides with conscious access (Dehaene & Changeux 2011). Hence, disentangling bottom-up and top-down multisensory processes is likely to be

an important step towards understanding the intricate links between multisensory integration and consciousness (De Meo et al. 2015).

Section 3: Which experimental approaches are most promising to characterize multisensory awareness?

Over the past decade, a growing number of studies have focused on the emergence of perceptual awareness in multisensory contexts. The majority of those studies have investigated how a signal arising from another sensory modality can modulate the access to visual awareness by using experimental paradigms in which visual stimuli, albeit presented on the retina, are suppressed from visual awareness using a variety of experimental paradigms such as attentional blink, masking, and multistable perception. Here we will focus on experimental approaches using *bistable visual stimuli* (Blake & Logothetis 2002) to investigate multisensory interactions during different states of visual awareness.

Bistable perception of ambiguous figures

Our visual system is often faced with perceptual ambiguity and perceptual decisions need to be made to efficiently interact with the external world. According to the Bayesian theory of perception (for review see Knill & Pouget 2004), the brain deals with perceptual uncertainty and ambiguity by representing sensory information in the form of probability distributions. If different perceptual interpretations have the same likelihood and are mutually exclusive, the visual system cannot "decide" in favour of one or the other and visual perception continuously oscillates between the two alternatives, a phenomenon called *bistable perception* (Dayan 1998). Bistable perception is thought to be generated by the competition between neural populations representing different interpretations of a visual stimulus (Blake

& Logothetis 2002). Perceptual bistability can arise from different forms of ambiguity: ambiguity in depth (e.g., the Necker cube, Necker, 1832), ambiguity in figure-ground segregation (e.g. Rubing's face-vase illusion, Rubin, 1915), ambiguity between high-level interpretations of images (e.g., Boring's young girl/old woman figure, Boring, 1930), ambiguity in the direction of motion (e.g., the kinetic depth effect, Doner et al. 1984). Investigating whether a signal arising from another sensory modality can disambiguate bistable perception favouring the access to awareness of the interpretation of the visual stimulus congruent with the cross-modal stimulus is an interesting approach to studying multisensory awareness. However, as pointed out by Deroy et al. (2014), since the bistable perception of ambiguous figures is to some extent under attentional control (Gómez et al. 1995; Horlitz & O'Leary 1993; Liebert & Burk 1985), it is difficult to disentangle the contribution of attention in mediating the effect of cross-modal stimulation on ambiguous figure perception. In fact, the interaction between bistable perception and tactile (visuo-tactile kinetic depth effect, Blake et al. 2004 and visuo-tactile Necker cube, Bruno et al. 2007), auditory (Robin's face/vase illusion with faces and voice uttering a syllable, Munhall et al. 2009) and olfactory (ambiguous motion direction associate with a particular smell, Kuang & Zhang 2014) stimuli depends on awareness of the congruent interpretation of the visual stimulus, namely, cross-modal stimulation only interacts with the representation of the stimulus dominating observer's perception, prolonging its duration.

Binocular Rivalry

A special case of perceptual bistability is *binocular rivalry* (Levelt 1965), that is caused by a conflict between monocular images rather than between different interpretations of the same monocular image. In a binocular rivalry display, incompatible images are contemporaneously presented to each eye (Fig. 1A), in this condition the two monocular images do not fuse into a coherent percept, but engage a strong competition for visual awareness that, in a "winner-

takes-all" dynamic, leads to perceptual oscillations between the two images: observer's perception is dominated by the stimulus presented to one eye for a few seconds until a perceptual switch occurs in favour the previously suppressed image (Fig. 1B). Importantly, during binocular rivalry, the suppressed visual stimulus is rendered invisible by the dominant one despite its presence on the retina. Compared to other forms of bistable perception, binocular rivalry is thought to be more automatic and stimulus-driven, ambiguous figures being more likely to be controlled by cognitive factors such as voluntary control or attention (Meng & Tong 2004). Importantly, voluntary attentional control over binocular rivalry is limited to dominance of the attended visual stimulus as the observer cannot voluntarily provoke a switch but only hold the dominant stimulus for a longer time (for a review on attention and binocular rivalry see Paffen & Alais, 2011). Even though experimental evidence shows that visual stimuli rendered invisible can exogenously capture attention and provide a cue for different kinds of visual tasks (Astle et al. 2010; Hsieh & Colas 2012; Hsieh et al. 2011; Lamy et al. 2015; Zhang & Fang 2012), it is more difficult to voluntarily select the suppressed visual stimulus during binocular rivalry. In this vein, if cross-modal stimulation influences the dynamics of binocular rivalry only by prolonging dominance durations of the congruent visual stimulus (Fig. 1C reports an example of visuo-haptic interactions during binocular rivalry), the effect could in principle be mediated by a cross-modal shift of attention or a higher-level cognitive decision. If, on the other hand, cross-modal stimulation shortens the suppression of the congruent visual stimulus (provoking a switch when the visual and cross-modal stimulus are incongruent, Fig. 1D), promoting the access to awareness of the suppressed visual stimulus, the effect is likely to reflect a genuine case of multisensory awareness. It has been argued (Deroy et al. 2014) that the cross-modal modulation of visual awareness does not represent multisensory awareness, but only a case of multisensory interaction on the basis that this experimental approach studies awareness in a unisensory

framework (for example the access to visual awareness) and not the establishment of multisensory awareness from different sensory modalities information (Deroy et al. 2014). However, in this case, the observer is not aware of a visual stimulus on its own, and awareness is built by integrating signals from different modalities and therefore we may potentially be able to consider it a case of multisensory awareness.

Figure 1 here

Several studies have reported multisensory effects on binocular rivalry depending on awareness, and therefore possibly mediated by attention: dominance durations of the congruent visual stimulus are prolonged by auditory (Conrad et al. 2010; Guzman-Martinez et al. 2012; Kang & Blake 2005; Lee et al. 2015) and nostril-specific olfactory stimulation (Zhou et al. 2010) and by imitation of a grasping movement rivalling against a checkerboard (Di Pace & Saracini 2014). A strict link between cross-modal attention and binocular rivalry has been demonstrated by a study showing that cross-modal stimulation enhances people's attentional control over binocular rivalry (van Ee et al. 2009). In this study, observers were asked to attend selectively to one of the rivalrous visual stimuli (which prolonged dominance durations of the attended stimulus compared to passive viewing), if either a sound or a vibration congruent with the attended visual stimulus was delivered simultaneously, dominance durations of the attended visual stimulus increased compared to the visual-only condition (van Ee et al. 2009). This result has recently been replicated using auditory and visual speech stimuli (Vidal & Barrès 2014): The latter researchers have shown that auditory syllables increase voluntary control over the rivalrous image of lips uttering the congruent syllable.

What about cross-modal stimuli interacting with the suppression of the congruent visual stimulus? Numerous experiments have demonstrated that haptic and auditory stimulation

interact with binocular rivalry by rescuing the congruent visual stimulus from binocular suppression (Conrad et al. 2010; Lunghi et al. 2010). In a first study, Lunghi et al. (2010), demonstrated that, during binocular rivalry between orthogonally oriented visual gratings, active exploration of a haptic grating, promoted dominance of the rivalrous visual grating congruent in orientation both by prolonging its dominance durations (delaying the time of a perceptual switch during congruent visuo-haptic stimulation) and by shortening its suppression (hastening the time of a perceptual switch during incongruent visuo-haptic stimulation) as compared to visual-only stimulation (Lunghi et al. 2010). The effect of haptic stimulation on the suppressed visual stimulus has been shown to depend critically on the match between visuo-haptic spatial frequencies (Lunghi et al. 2010) and orientations (Lunghi & Alais 2013) and on the collocation of the haptic and visual stimuli (Lunghi & Morrone 2013), indicating that the visual and haptic stimuli have to be perceived as being part of the same object and not simply cognitively associated. Moreover, a binocular rivalry experiment investigating suppression depth (the difference between contrast detection thresholds measured during dominance and suppression phases of binocular rivalry) during haptic stimulation (Lunghi & Alais 2015) has shown that haptic stimulation influences the dynamics of binocular rivalry mainly by preventing the congruent visual stimulus from becoming deeply suppressed (Lunghi & Alais 2015). This study clarifies a possible confound: during binocular rivalry, the monocular signals mutually inhibit each other, thus, in principle, touch could shorten suppression of the congruent visual stimulus both by interacting with it or potentially by interfering with the incongruent dominant stimulus, reducing its strength. By demonstrating that congruent touch improves contrast detection thresholds during suppression and incongruent touch does not have a masking effect on contrast detection thresholds during dominance (i.e., contrast discrimination thresholds are no higher during

incongruent touch), Lunghi & Alais (2015) have demonstrated that cross-modal stimulation during binocular rivalry actually boosts the suppressed visual signal.

Similar effects on the suppressed visual stimulus have been reported for voluntary action (voluntarily controlling the motion direction of one of the rivalring stimuli by an active movement of the arm shortens its suppression, Maruya et al. 2007), simple and naturalistic motion sounds (Blake et al. 2004; Conrad et al. 2010; Conrad et al. 2013), olfaction (suppression duration of either the picture of a marker or a rose is reduced when a congruent odorant is smelled, Zhou et al., 2010), ecologically relevant sounds (hearing a bird singing reduces suppression of the picture of a bird, Chen et al. 2011) and temporal events (auditory and tactile temporal events combine to synchronize binocular rivalry between visual stimuli differing in temporal frequency, Lunghi et al. 2014).

Continuous Flash Suppression

In order to selectively study the effect of cross-modal stimulation on visual stimuli undergoing interocular suppression the method of *continuous flash suppression* (CFS) can be used, for it allows deep and constant suppression of a salient visual stimulus over extended periods of time (Tsuchiya & Koch 2005). When one eye is continuously flashed with different, contour-rich, high-contrast random patterns (e.g. white noise, Mondrian patterns, scrambled images) at about 10 Hz, information presented to the other eye is perceptually suppressed for extended periods of time (up to 3 minutes or more). Suppression provoked by continuous flashes has been shown to summate, resulting not only in longer suppression periods, but also in deeper suppression of the other eye: detection thresholds of probes presented to the suppressed eye during CFS are in fact elevated of a 20-fold factor compared to monocular viewing, compared with a 3-fold elevation observed during binocular rivalry

(Tsuchiya et al. 2006). Importantly, in binocular rivalry perception continuously alternates between the monocular images leading to some cognitive awareness about the suppressed stimulus, for it was the dominant one before the perceptual switch. By contrast, during CFS the coherent stimulus is deeply suppressed the flashing masks, so the observer is totally unaware of the suppressed visual stimulus, not only at the perceptual (the stimulus is invisible), but also at the cognitive level (no information about the suppressed visual stimulus is available to the observer either from memory or predictions). If a visual stimulus were released from CFS by a congruent cross-modal stimulus gaining access to visual awareness it would provide a case of multisensory awareness, or at least of awareness that has been induced multisensorially.

Recent evidence has described cross-modal influences on CFS. Alsius & Munhall (2013) have shown that the movie of lips uttering a sentence made invisible by CFS is rescued from suppression earlier if observers listen to a voice speaking the sentence uttered by the movie as compared to an incongruent sentence. Salomon et al. (2013) have reported a similar result for proprioceptive signals, demonstrating that the image of a hand (perceptually projected on the observer's real hand) is suppressed for a shorter time during CFS if it matches the position of the observer's own hand as compared to an incongruent position (Salomon et al. 2013). A recent study from the same group (Salomon et al. 2015) has also shown a facilitation for congruent versus incongruent visuo-vestibular stimulation during CFS (Salomon et al. 2015). Finally, auditory facilitation of suppressed visual stimuli has been shown to depend on spatial collocation between the cross-modal stimuli both along the azimuth (Aller et al. 2015) and depth planes (Yang & Yeh 2014).

Taken together, the results reviewed here suggest that binocular rivalry and CFS are two promising techniques for characterizing multisensory awareness: first, suppressed visual stimuli are boosted into visual awareness via very specific mechanisms that rely on classical multisensory congruency cues that indicate whether sensory signals are caused by a common event; second binocular rivalry suppression and continuous flash suppression are impenetrable to voluntary attention (indicating a genuine multisensory effect is unlikely mediated by cross-modally driven attentional shifts); third a variety of sensory signals contribute to the multisensory enhancement of awareness (audition, touch, proprioception, voluntary action, olfaction, and the vestibular system), indicating a real supramodal mechanism mediating and consolidating awareness. Interestingly, a recent study (Di Luca et al. 2010) has shown that observers can learn to use invisible information (for example a vertical disparity gradient masked by other visual stimuli) to disambiguate visual perception in a bistable display (Di Luca et al. 2010), this suggests that similar learning paradigms using subliminal cross-modal stimuli in combination with either binocular rivalry or CFS could be used in the future to study cross-modal awareness.

Concluding remarks

Our discussion has highlighted substantial advances in our understanding of multisensory awareness over the past decade. Nevertheless, research into the relationship between multisensory integration and perceptual awareness faces a couple of unresolved challenges:

First, it is controversially debated which perceptual experiences are necessarily associated with multisensory awareness. In the face of uncertainty concerning the underlying causal structure of the world the brain does often not integrate sensory signals into one unified

multisensory percept. For instance, in the spatial ventriloquist illusion participants tend to report different locations for the visual and the auditory signal sources with the perceived sound location being shifted towards the visual signal and the visual location towards the auditory signal depending on the relative reliabilities (Körding et al. 2007; Rohe & Noppeney 2015, 2016).

Can a sound percept that is influenced by a visual signal be considered an example of multisensory awareness? Further, when participants report both, the perceived sound and the perceived visual location, are they concurrently aware of both signals or do they rapidly switch their attention and awareness to perceptual and memory representations from different sensory modalities? Finally, in cases where participants are thought to integrate sensory signals into one unified percept and report identical locations for both sensory signals, does this guarantee the emergence of integrated multisensory awareness or are participants simply not able to dissociate between the two sorts of unisensory awareness in their report? In the light of these puzzling questions, it is interesting to note that participants were not able to perceive and report the motion direction both in vision and touch when being presented concurrently with a bistable motion quartet in the visual and tactile modalities (Conrad et al., 2012). Thus, at least in situations where perception in the individual sensory modalities requires sustained temporal perceptual binding (such as in the case of apparent motion), multisensory awareness may not necessarily emerge, instead in these instances awareness switches between sensory modalities such as vision and touch.

Second, numerous studies in anaesthetized animals have demonstrated that multisensory interactions can emerge in the absence of awareness (Stein & Meredith 1993). Yet, their relevance for conscious perception remains to be determined. While accumulating evidence suggests that aware signals can boost unaware signals into awareness, little is known about whether the reverse is also true. Can unaware signals in one sensory modality influence

perception in another sensory modality? Experiments focusing on the latter are more informative, because the former could simply be explained by non-specific top-down effects. Moreover, experiments may subliminally present signals in two sensory modalities that can be integrated into a unified percept to show that the subliminally integrated estimate influences subsequent conscious perception.

Third, research into perceptual awareness in unisensory contexts has recently re-focused on classical metacognitive questions and asked to which extent participants can recognize their perceptual performance and abilities. This is an exciting as yet little explored avenue that would provide further insights into the emergence of multisensory integration, perception and awareness.

Acknowledgements:

This work was supported by EU FP7 Framework Programme FPT/2007-2013, ECSPLAIN, n.338866 (CL), European Research Council ERC-STG multsens (UN), the AHRC Rethinking the Senses grant AH/L007053/1 (OD & CS) and a Marie Skłodowska-Curie Fellowship (NF).

References:

Adam, R. & Noppeney, U., 2014. A phonologically congruent sound boosts a visual target into perceptual awareness. *Frontiers in Integrative Neuroscience*, 8(September), p.70.

Alais, D. & Burr, D., 2004. Ventriloquist effect results from rear-optimal bimodal integration. *Current Biology*, 14(3), pp.257–262.

Allen, J., Kraus, N. & Bradlow, A., 2000. Neural representation of consciously imperceptible speech sound differences. *Perception & Psychophysics*, 62(7), pp.1383–1393.

Aller, M., Giani, A., Conrad, V., Watanabe, M. & Noppeney, U., 2015. A spatially collocated sound thrusts a flash into awareness. *Frontiers in Integrative Neuroscience*, 9(February), pp.1–8.

Alsius, A. & Munhall, K.G., 2013. Detection of audiovisual speech correspondences without visual awareness. *Psychological Science*, 24(4), pp.423–431.

Andersen, T.S., Tiippana, K. & Sams, M., 2004. Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, 21(3), pp.301–308.

Arzi, A., Shedlesky, L., Ben-Shaul, M., Nasser, K., Oksenberg, A., Hairston, I.S. & Sobel, N., 2012. Humans can learn new information during sleep. *Nature Neuroscience*, 15(10), pp.1460–1465.

Astle, D.E., Nobre, A.C. & Scerif, G., 2010. Subliminally presented and stored objects capture spatial attention. *The Journal of Neuroscience*, 30(10), pp.3567–3571.

Balduzzi, D. & Tononi, G., 2008. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Computational Biology*, 4(6), p.e1000091.

Bekinschtein, T.A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L. & Naccache, L., 2009. Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), pp.1672–1677.

Bertelson, P. & Aschersleben, G., 1998. Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review*, 5(3), pp.482–489.

Blake, R. & Logothetis, N.K., 2002. Visual competition. *Nature Reviews. Neuroscience*, 3(1), pp.13–21.

Blake, R., Sobel, K. V. & James, T.W., 2004. Neural synergy between kinetic vision and touch. *Psychological Science*, 15(6), pp.397–402.

Blanke, O., 2012. Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews. Neuroscience*, 13(8), pp.556–571.

Boring, E.G., 1930. A new ambiguous figure. *The American Journal of Psychology*, 42(3), p.444.

Bruno, N., Jacomuzzi, A., Bertamini, M. & Meyer, G., 2007. A visual-haptic Necker cube reveals temporal constraints on intersensory merging during perceptual exploration. *Neuropsychologia*, 45(3), pp.469–475.

Chalmers, D. & Bayne, T., 2003. What is the unity of consciousness? In A. Cleeremans & C. Frith, eds. *The Unity of Consciousness: Binding, Integration, and Dissociation*. Oxford University Press.

Chen, Y.C. & Spence, C., 2010. When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, 114(3), pp.389–404.

Chen, Y.-C. & Spence, C., 2011a. Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), pp.1554–1568.

Chen, Y.-C. & Spence, C., 2011b. The crossmodal facilitation of visual object representations by sound: evidence from the backward masking paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6), pp.1784–1802.

Chen, Y.-C., Yeh, S.-L. & Spence, C., 2011. Crossmodal constraints on human perceptual awareness: auditory semantic modulation of binocular rivalry. *Frontiers in Psychology*, 2(SEP), p.212.

Conrad, V., Bartels, A., Kleiner, M. & Noppeney, U., 2010. Audiovisual interactions in binocular rivalry. *Journal of Vision*, 10(10), p.27.

Conrad, V., Kleiner, M., Bartels, A., Hartcher O'Brien, J., Bülthoff, H.H. & Noppeney, U., 2013. Naturalistic stimulus structure determines the integration of audiovisual looming signals in binocular rivalry. *PLoS ONE*, 8(8), p.e70710.

Conrad, V., Vitello, M.P. & Noppeney, U., 2012. Interactions between apparent motion rivalry in vision and touch. *Psychological Science*, 23(8), pp.940–948.

Critchley, H.D. & Harrison, N.A., 2013. Visceral influences on brain and behavior. *Neuron*, 77(4), pp.624–638.

Dayan, P., 1998. A hierarchical model of binocular rivalry. *Neural Computation*, 10(5), pp.1119–1135.

Dehaene, S., 2001. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1-2), pp.1–37.

Dehaene, S. & Changeux, J.-P., 2011. Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), pp.200–227.

Deroy, O., 2014. The unity assumption and the many unities of consciousness. In D. Bennett & C. Hill, eds. *Sensory Integration and the Unity of Consciousness.* Cambridge, MA, US: MIT Press, p. 105.

Deroy, O., Chen, Y. & Spence, C., 2014. Multisensory constraints on awareness. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 369(1641), p.20130207.

Doner, J., Lappin, J.S. & Perfetto, G., 1984. Detection of three-dimensional structure in moving optical patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 10(1), pp.1–11.

van Ee, R., van Boxtel, J.J. a, Parker, A.L. & Alais, D., 2009. Multisensory congruency as a mechanism for attentional control over perceptual selection. *Journal of Neuroscience*, 29(37), pp.11641–11649.

Ernst, M.O. & Banks, M.S., 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), pp.429–433.

Faivre, N., Mudrik, L., Schwartz, N. & Koch, C., 2014. Multisensory integration in complete unawareness: Evidence from audiovisual congruency priming. *Psychological Science*, 25(11), pp.2006–2016.

Faivre, N., Salomon, R. & Blanke, O., 2015. Visual consciousness and bodily self-consciousness. *Current Opinion in Neurobiology*, 28(1), pp.23–28.

Gallace, A. & Spence, C., 2014. *In touch with the future*, Oxford University Press.

Gallace, A. & Spence, C., 2008. The cognitive and neural correlates of "tactile consciousness": A multisensory perspective. *Consciousness and Cognition*, 17(1), pp.370–407.

Gau, R. & Noppeney, U., 2016. How prior expectations shape multisensory perception. *NeuroImage*, 124(Pt A), pp.876–886.

Ghazanfar, A. a & Schroeder, C.E., 2006. Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10(6), pp.278–285.

Ghose, G.M. & Maunsell, J., 1999. Specialized representations in visual cortex: a role for binding? *Neuron*, 24(1), pp.79–85.

Giani, A.S., Belardinelli, P., Ortiz, E., Kleiner, M. & Noppeney, U., 2015. Detecting tones in complex auditory scenes. *NeuroImage*, 122, pp.203–213.

Gómez, C., Argandoña, E.D., Solier, R.G., Angulo, J.C. & Vázquez, M., 1995. Timing and competition in networks representing ambiguous figures. *Brain and Cognition*, 29(2), pp.103–114.

De Graaf, T.A., Hsieh, P.J. & Sack, A.T., 2012. The "correlates" in neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, 36(1), pp.191–197.

Gutschalk, A., Micheyl, C. & Oxenham, A.J., 2008. Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology*, 6(6), pp.1156–1165.

Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J. & Suzuki, S., 2012. Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology*, 22(5), pp.383–388.

Haynes, J.D., Driver, J. & Rees, G., 2005. Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron*, 46(5), pp.811–821.

Heffner, R.S. & Heffner, H.E., 1992a. Evolution of sound localization in mammals. In D. B. Webster, A. N. Popper, & R. R. Fay, eds. *The Evolutionary Biology of Hearing*. Springer Verlag, pp. 691–715.

Heffner, R.S. & Heffner, H.E., 1992b. Visual factors in sound localization in mammals. *The*

Journal of Comparative Neurology, 317(3), pp.219–232.

Hillis, J.M., Ernst, M.O., Banks, M.S. & Landy, M.S., 2002. Combining sensory information: mandatory fusion within, but not between, senses. *Science*, 298(5598), pp.1627–1630.

Horlitz, K.L. & O'Leary, A., 1993. Satiation or availability? Effects of attention, memory, and imagery on the perception of ambiguous figures. *Perception & Psychophysics*, 53(6), pp.668–681.

Hsiao, J.Y., Chen, Y.C., Spence, C. & Yeh, S.L., 2012. Assessing the effects of audiovisual semantic congruency on the perception of a bistable figure. *Consciousness and Cognition*, 21(2), pp.775–787.

Hsieh, P.-J. & Colas, J.T., 2012. Awareness is necessary for extracting patterns in working memory but not for directing spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), pp.1085–1090.

Hsieh, P.-J., Colas, J.T. & Kanwisher, N., 2011. Pop-out without awareness: Unseen feature singletons capture attention only when top-down attention is available. *Psychological Sciencece*, 22(9), pp.1220–1226.

Jackson, C. V., 1953. Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology*, 5(2), pp.52–65.

Kang, M.-S. & Blake, R., 2005. Perceptual synergy between seeing and hearing revealed during binocular rivalry. *Psychologija*, 32.

Kayser, C., Logothetis, N.K. & Panzeri, S., 2010. Visual enhancement of the information representation in auditory cortex. *Current Biology*, 20(1), pp.19–24.

Knill, D.C. & Pouget, A., 2004. The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), pp.712–719.

Körding, K.P., Beierholm, U., Ma, W.J., Quartz, S., Tenenbaum, J.B. & Shams, L., 2007. Causal inference in multisensory perception. *PLoS ONE*, 2(9), p.e943.

Kruger, H., Collins, T. & Cavanagh, P., 2014. Similar effects of saccades on auditory and visual localization suggest common spatial map. *Journal of Vision*, 14(10), pp.1232–1232.

Kuang, S. & Zhang, T., 2014. Smelling directions: Olfaction modulates ambiguous visual motion perception. *Scientific Reports*, 4, p.5796.

Lamme, V.A.F., 2006. Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), pp.494–501.

Lamme, V.A.F. & Roelfsema, P.R., 2000. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23(11), pp.571–579.

Lamy, D., Alon, L., Carmel, T. & Shalev, N., 2015. The role of conscious perception in attentional capture and object-file updating. *Psychological Science*, 26(1), pp.48–57.

Lee, H. & Noppeney, U., 2011. Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proceedings of the National Academy of Sciences of*

the United States of America, 108(51), pp.E1441–E1450.

Lee, H. & Noppeney, U., 2014. Temporal prediction errors in visual and auditory cortices. *Current Biology*, 24(8), pp.R309–R310.

Lee, M., Blake, R., Kim, S. & Kim, C.-Y., 2015. Melodic sound enhances visual awareness of congruent musical notes, but only if you can read music. *Proceedings of the National Academy of Sciences of the United States of America*, 112(27), pp.8493–8498.

Levelt, W.J., 1965. *On binocular rivalry*, Soesterberg, Netherlands: Institution for Perception.

Liang, M., Mouraux, A., Hu, L. & Iannetti, G.D., 2013. Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nature Communications*, 4(May), p.1979.

Liebert, R.M. & Burk, B., 1985. Voluntary control of reversible figures. *Perceptual and Motor Skills*, 61(3 Pt 2), pp.1307–1310.

Di Luca, M., Ernst, M.O. & Backus, B.T., 2010. Learning to use an invisible visual signal for perception. *Current Biology*, 20(20), pp.1860–1863.

Lunghi, C. & Alais, D., 2015. Congruent tactile stimulation reduces the strength of visual suppression during binocular rivalry. *Scientific Reports*, 5, p.9413.

Lunghi, C. & Alais, D., 2013. Touch interacts with vision during binocular rivalry with a tight orientation tuning. *PLoS ONE*, 8(3), p.e58754.

Lunghi, C., Binda, P. & Morrone, M.C., 2010. Touch disambiguates rivalrous perception at early stages of visual analysis. *Current Biology*, 20(4), pp.R143–R144.

Lunghi, C. & Morrone, M.C., 2013. Early interaction between vision and touch during binocular rivalry. *Multisensory Research*, pp.1–16.

Lunghi, C., Morrone, M.C. & Alais, D., 2014. Auditory and tactile signals combine to influence vision during binocular rivalry. *The Journal of Neuroscience*, 34(3), pp.784–792.

Maeda, F., Kanai, R. & Shimojo, S., 2004. Changing pitch induced visual motion illusion. *Current Biology*, 14(23), pp.R990–R991.

Maruya, K., Yang, E. & Blake, R., 2007. Voluntary action influences visual competition. *Psychological Science*, 18(12), pp.1090–1098.

McGurk, H. & MacDonald, J., 1976. Hearing lips and seeing voices. *Nature*, 264(5588), pp.746–748.

Meng, M. & Tong, F., 2004. Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, 4(7), pp.539–551.

De Meo, R., Murray, M.M., Clarke, S. & Matusz, P.J., 2015. Top-down control and early multisensory processes: chicken vs. egg. *Frontiers in Integrative Neuroscience*, 9, p.17.

Meredith, M.A. & Stein, B.E., 1986. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 56(3), pp.640–662.

Meyer, K., Kaplan, J.T., Essex, R., Webber, C., Damasio, H. & Damasio, A., 2010. Predicting visual stimuli on the basis of activity in auditory cortices. *Nature Neuroscience*, 13(6), pp.667–668.

Mudrik, L., Faivre, N. & Koch, C., 2014. Information integration without awareness. *Trends in Cognitive Sciences*, 18(9), pp.488–496.

Munhall, K.G., Gribble, P., Sacco, L. & Ward, M., 1996. Temporal constraints on the McGurk effect. *Perception & Psychophysics*, 58(3), pp.351–362.

Munhall, K.G., ten Hove, M.W., Brammer, M. & Paré, M., 2009. Audiovisual integration of speech in a bistable illusion. *Current Biology*, 19(9), pp.735–739.

Necker, L.A., 1832. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid. *Philosophical Magazine Series 3*, 1(5), pp.329–337.

Noel, J.-P., Wallace, M. & Blake, R., 2015. Cognitive neuroscience: Integration of sight and sound outside of awareness? *Current Biology*, 25(4), pp.R157–R159.

Olivers, C.N.L. & Van der Burg, E., 2008. Bleeping you out of the blink: sound saves vision from oblivion. *Brain Research*, 1242(2000), pp.191–199.

Di Pace, E. & Saracini, C., 2014. Action imitation changes perceptual alternations in binocular rivalry. *PLoS ONE*, 9(5), p.e98305.

Paffen, C.L.E. & Alais, D., 2011. Attentional modulation of binocular rivalry. *Frontiers in Human Neuroscience*, 5, p.105.

Palmer, T.D. & Ramsey, A.K., 2012. The function of consciousness in multisensory integration. *Cognition*, 125(3), pp.353–364.

Partan, S. & Marler, P., 1999. Communication goes multimodal. *Science*, 283(5406), pp.1272–1273.

Ro, T., Breitmeyer, B., Burton, P., Singhal, N.S. & Lane, D., 2003. Feedback contributions to visual awareness in human occipital cortex. *Current Biology*, 13(12), pp.1038–1041.

Rohe, T. & Noppeney, U., 2015. Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biology*, 13(2), p.e1002073.

Rohe, T. & Noppeney, U., 2016. Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Current Biology*, 1, pp.1–6.

Romei, V., Murray, M.M., Cappe, C. & Thut, G., 2009. Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology*, 19(21), pp.1799–1805.

Roskies, A.L., 1999. The binding problem. *Neuron*, 24(1), pp.7–9.

Rubin, E., 1915. Synsoplevede figurer (Visually experienced figures). In *Studier i Psykologisk Analyse*. Copenhagen: Gyldendal.

von Saldern, S. & Noppeney, U., 2013. Sensory and striatal areas integrate auditory and visual signals into behavioral benefits during motion discrimination. *Journal of*

Neuroscience, 33(20), pp.8841–8849.

Salomon, R., Kaliuzhna, M., Herbelin, B. & Blanke, O., 2015. Balancing awareness: Vestibular signals modulate visual consciousness in the absence of awareness. *Consciousness and Cognition*, 36, pp.289–297.

Salomon, R., Lim, M., Herbelin, B., Hesselmann, G. & Blanke, O., 2013. Posing for awareness: proprioception modulates access to visual consciousness in a continuous flash suppression task. *Journal of Vision*, 13(7), p.2.

Schroeder, C.E. & Foxe, J., 2005. Multisensory contributions to low-level, "unisensory" processing. *Current Opinion in Neurobiology*, 15(4), pp.454–458.

Shams, L., Kamitani, Y. & Shimojo, S., 2000. Illusions. What you see is what you hear. *Nature*, 408(6814), p.788.

Spence, C. & Bayne, T., 2015. Is consciousness multisensory? In D. Stokes, M. Matthen, & S. Biggs, eds. *Perception and Its Modalities*. Oxford, UK: Oxford University Press, pp. 95–132.

Spence, C., Smith, B. & Auvray, M., 2015. Confusing tastes and flavours. In D. Stokes, M. Matthen, & S. Biggs, eds. *Perception and Its Modalities*. Oxford,UK: Oxford University Press, pp. 247–274.

Stanford, T.R., 2005. Evaluating the operations underlying multisensory integration in the cat superior colliculus. *Journal of Neuroscience*, 25(28), pp.6499–6508.

Stein, B.E. & Meredith, M.A., 1993. *The merging of the senses*, Cambridge, MA, US: The MIT Press.

Stein, B.E. & Stanford, T.R., 2008. Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews. Neuroscience*, 9(4), pp.255–266.

Stein, B.E., Stanford, T.R., Ramachandran, R., Perrault, T.J. & Rowland, B.A., 2009. Challenges in quantifying multisensory integration: Alternative criteria, models, and inverse effectiveness. *Experimental Brain Research*, 198(2-3), pp.113–126.

Stevenson, R.J. & Attuquayefio, T., 2013. Human olfactory consciousness and cognition: its unusual features may not result from unusual functions but from limited neocortical processing resources. *Frontiers in Psychology*, 4(NOV), p.819.

Tsuchiya, N. & Koch, C., 2005. Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, 8(8), pp.1096–1101.

Tsuchiya, N., Koch, C., Gilroy, L. a & Blake, R., 2006. Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of Vision*, 6(10), pp.1068–1078.

Vetter, P., Smith, F.W. & Muckli, L., 2014. Decoding sound and imagery content in early visual cortex. *Current Biology*, 24(11), pp.1256–1262.

Vidal, M. & Barrès, V., 2014. Hearing (rivaling) lips and seeing voices: how audiovisual interactions modulate perceptual stabilization in binocular rivalry. *Frontiers in Human Neuroscience*, 8, p.677.

Vroomen, J. & de Gelder, B., 2004. Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. In G. Calvert, C. Spence, & B. E. Stein, eds. *The Handbook of Multisensory Processes*. MIT Press, MA, USA, pp. 141–146.

Werner, S. & Noppeney, U., 2010a. Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *The Journal of neuroscience*, 30(7), pp.2662–2675.

Werner, S. & Noppeney, U., 2010b. Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cerebral Cortex*, 20(8), pp.1829–1842.

Wolfe, J.M. & Cave, K.R., 1999. The psychophysical evidence for a binding problem in human vision. *Neuron*, 24(1), pp.11–17.

Yang, Y.H. & Yeh, S.L., 2014. Unmasking the dichoptic mask by sound: Spatial congruency matters. *Experimental Brain Research*, 232(4), pp.1109–1116.

Zhang, X. & Fang, F., 2012. Object-based attention guided by an invisible object. *Experimental Brain Research*, 223(3), pp.397–404.

Zhou, W., Jiang, Y., He, S. & Chen, D., 2010. Olfaction modulates visual perception in binocular rivalry. *Current Biology*, 20(15), pp.1356–1358.

Figures:

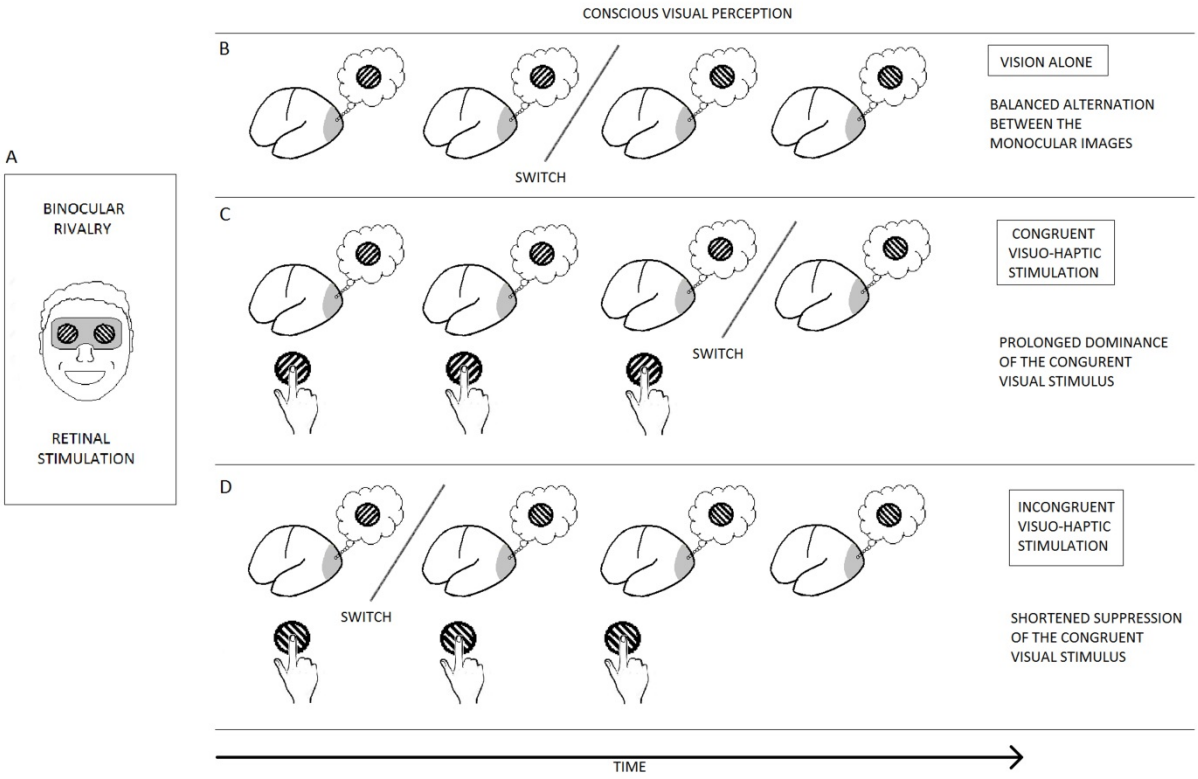

Figure 1. Diagram of a binocular rivalry display and possible effects of cross-modal stimulation on rivalrous visual perception. (A) An example of dichoptic stimulation in which orthogonal gratings are separately presented to the eyes, the resulting conscious perception (B) is dominated by one of the two monocular images until a perceptual switch occurs in favour of the other visual stimulus. Normally dominance duration of the rivalrous stimuli is balanced. (C) Example of cross-modal stimulation prolonging dominance of the congruent visual stimulus during binocular rivalry: if the observer touches a haptic grating parallel to the visual grating dominating rivalrous perception, the switch towards the orthogonal (incongruent) visual grating is delayed as compared to visual-only stimulation. (D) Example of cross-modal stimulation shortening the suppression of the congruent visual stimulus during binocular rivalry: if the observer touches a haptic grating orthogonal to the visual grating dominating rivalrous perception, the switch towards the parallel (congruent) visual grating occurs earlier compared to visual-only stimulation.